# COMMERCIAL PAYROLL SOFTWARE AND THE QUALITY OF EMPLOYMENT DATA

**Karen L. Goldenberg, Susan E. Moore, Richard J. Rosen, U.S. Bureau of Labor Statistics**
**Susan Moore, BLS, Suite 4840, 2 Massachusetts Avenue, N.E., Washington, DC 20212**

**KEY WORDS: Measurement error; Response error; Establishment data; Payroll records.**

## I. Introduction and Background

The Bureau of Labor Statistics (BLS) measures monthly job fluctuations by tabulating employee counts supplied to two different statistical programs: the Covered Employment and Wages Program (ES-202), which receives data from State Unemployment Insurance (UI) systems, and the Current Employment Statistics (CES) Program. Employers submit employee counts for UI purposes on State Unemployment Insurance Quarterly Contribution Reports (QCRs), and to CES by means of the BLS monthly Report on Employment, Payroll, and Hours. Data for these employee counts frequently come from reports associated with payroll records. Since the 1980s, growing numbers of payrolls have been produced by standardized software (Werking et al., 1993), either supplied by commercial payroll processing firms (PPFs) or generated for the commercial market by payroll software developers (PSDs). The quality of employment data in both statistical programs is thus increasingly subject to the way PPFs and PSDs define and calculate employment.

The importance of the payroll processing industry came to light in June, 1992, following the release of a much larger-than-usual adjustment from the 1991 sample-based CES data to the universe-based ES-202 data. Research eventually attributed the situation to the payroll processing industry's introduction of improved reporting procedures early in 1991 (Werking et al., 1993). In an attempt to better understand the overall importance and influence of the industry on the quality of employment data, BLS launched the PPF Response Analysis Survey (RAS).

## II. General Survey Plan

The PPF RAS studied the payroll procedures and reporting practices used by the payroll processing industry to prepare data for UI purposes. (Preliminary study showed that PPFs did not usually prepare data for CES.) RAS results are helping BLS to understand the extent of the industry and the limitations of employment data prepared by it.

The PPF RAS was conducted in two distinct phases. The first phase was a pilot study and consisted of personal visit interviews with the larger firms in the payroll processing industry. These firms were known to provide services covering a substantial portion of U.S. employment. Twenty-four firms were targeted for interview, and 21 of these were actually interviewed by personal visit.

The next phase of the RAS involved telephone interviews with the remaining firms in the payroll processing industry.[1] There were two primary goals for the telephone portion of the RAS.

- To continue the research begun in the pilot study, ultimately establishing a reporting profile for the entire industry.

- To determine the magnitude and direction of response error in the data produced by the payroll processing industry, and to assess the impact of this error on employment data reported to the statistical programs.

The BLS Atlanta Data Collection Center (DCC) conducted the telephone interviews between June and September, 1993. Survey procedures included the use of advance letters, a Computer Assisted Telephone Interview (CATI) instrument, up to eight attempts to reach a sample unit, and thank you letters to respondents for completed interviews.

## III. Employment Definitions

The central focus for this study was the reporting of employment. The BLS definition of employment is **"the number of employees who worked during or received pay for the pay period including the 12th day of the calendar month."** Correct reporting depends on a number of factors, including conformity to this definition, availability of information, and correct use of data.

There are three separate components to the BLS definition of employment: method, timing, and content. If the reported employment count is to meet BLS definitions for the two programs, all three components must be handled correctly.

- The *method* by which the employment counts are derived can vary widely, and is probably determined by the content of the payroll system's standard outputs. The count most consistent with the BLS definition is *an unduplicated count of individuals* working or receiving a check or other form of payment, which prevents the double counting of employees receiving more that one check. However, reporters may use counts of active employees, employee records, employees who received checks, or counts of checks issued. Most of the incorrect reporting methods tend to overstate employment.

---

[1] The RAS was planned as a census of the entire industry. However, the number of firms in the payroll processing industry was larger than expected. Therefore, we selected a sample of about 250 firms.

- The *time period* for which employment is reported is the second dimension. The pay period including the 12th of the month is consistent with the BLS definition. Again, most incorrect reporting results in an overcount. Typical errors include reporting a count for the entire month, or reporting the total quarterly employment each month. In rare cases, undercounting results from using an incorrect time reference period, such as an employment figure that counts the number of employees on a specific day of the month.

- The *content* (i.e., who is counted) is the third component and can introduce another type of response error. This occurs if the concept of employment for CES and UI is misunderstood. Content error also tends to produce overcounts, because reporters are more likely to include employees who are not covered than to exclude covered employees. Of the three sources, content error has the smallest impact on overall counts.

Other variable aspects of employment reporting relate to how PPFs and payroll software products present the data to the user, as well as how the employer uses the data made available by the PPF. These variables clearly affect the quality of employment counts submitted.

## IV. The Payroll Processing Industry

This section describes the payroll processing industry and its contributions to monthly employment counts. We deal here with two types of businesses. **PPFs** collect information on hours worked, pay rates, deductions, etc. from their clients and use that information to generate paychecks and payroll reports. **PSDs** write and sell computer software to accept information on hours worked, pay rates, and deductions and to generate paychecks and payroll reports. Some firms are engaged in both payroll processing and software development. The industry also includes payroll tax filing services, businesses which accept data from clients and use those data to prepare and submit payroll taxes. The tax filing services do not generate employment counts, and we excluded them as out of scope.

This report emphasizes the effects of the payroll processing industry on employment counts compiled by BLS. The relationship between the payroll industry and those counts is not necessarily a direct one. Although many PPFs offer services that can provide a direct input to State employment reports, clients are not obligated to purchase those services. Payroll software systems and packages may contain functions that streamline reporting, if the licensee or purchaser chooses to use them. An employer who prepares his or her own UI documents may obtain data from a payroll report or develop the figures from another source. Even if the data have been computed correctly on the payroll report or other source document, the employer must find and extract the figures that correspond to BLS definitions.

For purposes of the analysis and discussion that follow, all results are presented as percentages weighted by probability of selection.[2] The resulting estimates are representative of the entire industry. When indicated, results for the PPFs may be shown as percentages weighted by the number of employees ultimately receiving services from those PPFs. The weighted data illustrate the effect of these firms on reported employment counts, and therefore the extent to which measurement error is a problem.

NOTE: Information on the distribution of the errors was subject to the extent of the knowledge of the respondents and was difficult to verify. In some cases, responses were inconsistent with other data collected. The team edited inconsistent responses if the data could not be verified.

## A. Descriptive Profile of Industry

For purposes of this survey, a firm provides ***payroll processing*** services if it collects basic wage and hour inputs each pay period, issues pay checks or direct deposits, and/or produces payroll summaries for clients. A firm supplies ***payroll software*** if it develops a software product that generates pay checks or payroll summaries, and sells, licenses, or leases that product directly to clients or end users.

While these definitions were helpful in identifying respondents, the pilot test showed that we can not draw a clear line between PPFs and PSDs. PPFs can and do sell, lease, or license their software as well as offer payroll processing services. In addition, a small number of PSDs have ventured into payroll processing as a business activity. These fuzzy boundaries between groups are important because the influence of the industry goes beyond that of its direct clients. The extent of that influence depends on the purchasers or licensees of the software and their size, information which is beyond the scope of this study.

Using the above definitions, we obtained data on 216 firms.[3] Of these, 44 provide only payroll processing services, 146 develop software and sell, lease, or license it, and 26 do both.

The questionnaire contained parallel sets of questions for PPFs and PSDs. When a firm supplied both types of services, we asked whether the company's primary business was payroll processing or software development. Two respondents said both areas were equal in importance, while the rest placed themselves into one category or the other. Based on these responses, and to avoid confusion in looking at the

---

[2] While PPFs were selected into the sample with certainty, some firms identified as PSDs were reclassified as PPFs based on their responses to a question on their primary activity.

[3] Actual interview counts are 19 from the pilot test, 7 from the pretest, and 190 from the DCC. We excluded 2 pilot study interviews with payroll tax filing services since they were out of scope.

effect of the industry on employment, we designated all companies as being either PPFs or PSDs, depending on their primary activity. For the companies that said that they were equally payroll processors and software developers, we separated PPF clients and covered employment from that of software users to avoid double counting the impact of these firms. However, we tabulated them in both places. As a result, the analysis is based on 218 companies, and describes the entire industry.[4]

Following these actions, the distribution by primary activity is 50 payroll processing firms and 168 payroll software developers.

### 1. PPF Coverage

Coverage is a measure of the extent to which the payroll processing industry influences overall employment counts compiled by BLS. For payroll processing firms, we looked at two different indicators of coverage: the number of client firms to which a PPF provides services, and the number of employees in those client companies. In addition, for both PPFs and PSDs, we looked at the concentration of clients in specific industries.

*Number of PPF client firms.* The RAS asked participating PPFs for the approximate number of client firms they supplied with payroll processing services. The majority of payroll processors each handle a small number of clients. Just under 19 percent of PPFs have fewer than 25 clients, and 52 percent have 250 or fewer clients. Seventeen percent report more than 2,500 clients, while only a very few firms (3 percent) handle more than 22,000.

*Number of employees in PPF client firms.* We estimated that approximately 24 million employees are covered by PPFs. Table 1 shows the distribution of this employment. Nearly 38 percent of PPFs provide payroll services covering no more than 5,000 employees, while another 45 percent of these processing companies cover from 5,000 to 100,000 employees. Only 17 percent of PPFs generate payrolls for more than 100,000 employees.

### 2. PSD Coverage

Coverage of payroll software was more difficult to determine. Most of the PSDs we visited leased or licensed their software, and so had information about their users. The PSDs we telephoned tended to sell their software, often at retail, and did not have this information.

The questionnaire for the telephone phase of the study asked about the number of copies of a product in use instead of the number of employees covered. In many cases, respondents were able to report only the total number of copies that had ever been sold.

We used this figure to arrive at an estimated 25 to 40 million employees covered by PSD products.[5]

Table 1.   PPFs by Client Firm Employment

| Approximate number of employees covered | Percent of PPFs |
|---|---|
| 1,000 or fewer | 18 |
| 1,001 to 5,000 | 20 |
| 5,001 to 10,000 | 14 |
| 10,001 to 50,000 | 17 |
| 50,001 to 100,000 | 14 |
| 100,001 to 300,000 | 12 |
| More than 300,000 | 5 |
| Total | 100 |

### 3. Length of pay period

The length of an employer's pay period is important for employment counts because those counts are supposed to refer to the pay period including the 12th of the month. However, that pay period can be anywhere from 7 to 31 days long, as most employers pay their employees weekly, biweekly, semi-monthly, or monthly. In addition, in some companies, different groups of employees are paid on different time schedules. One important question is whether the software permits the client or user to associate payrolls and numbers of employees with specific pay period dates. Table 2 shows how PPFs and PSDs maintain the length of pay period information.

Table 2.  Length of Pay Period in Payroll Software

| Length of pay period indicator | Percent PPFs | Percent PSDs |
|---|---|---|
| Start or end dates | 15 | 18 |
| Code for length of time period | 56 | 42 |
| Date and length of time period | 27 | 26 |
| Other (includes up to client) | 2 | 14 |

The fact that fewer than half of these payroll systems only have a length of pay period indicator rather than specific dates is of some concern to BLS. Although employers submit QCRs only once a quarter, the employment count is based on a single pay period each month. If the software does not identify start or end dates for the pay period, the proper reference period becomes more difficult to determine.[6]

### 4. Data Availability for Reporting Employment

*QCR.* The most important question is whether the software can supply the PPF, the client, or the system user with the employment counts needed to

---

[4] Several firms that provide both types of services do so in extremely unequal numbers. For example, one of the largest software firms had recently started a payroll processing business using its own software. The impact of this firm is generally through the sale of its software, not through the clients for which it prepares paychecks.

[5] The estimate ranges are based on an average of 10 employees and 20 employees per user company. The average size of a firm in the UI program is 20 employees.

[6] One possibility we did not explore is that the software may contain no length of pay period designator. A microcomputer payroll package purchased for use by one company doesn't really need one, since the user generates a payroll each pay period and knows exactly how long the pay period is.

prepare State and BLS forms. For the QCR, the answer is encouraging. Ninety-one percent of respondents said that their software produces the specific employment counts needed for the QCR, with no difference between PPFs and PSDs.[7] The percentage rose to 98 for PPFs and to 94 for PSDs when we asked if the system produced *any* reports that contained employee counts.

*CES.* During the interview we gave respondents the CES definition of employment and asked if they produced a specific count that conforms to it. If so, we asked PPF respondents if that count was made available each pay period, which it must be if respondents are to use the information for the CES.

Roughly 42 percent of the respondents said they produced the CES count, with no difference between PPFs and PSDs. Approximately 68 percent of the PPFs who produce the correct count also do so each pay period. If the count is available from a software product, the user can generate it each pay period.

5. Services Supporting Employment Counts

PPFs provide their clients with two important types of services that facilitate the submission of employment counts on State employment reports. These include filing a client's payroll taxes directly with the State, and preparing forms that contain employment counts. Payroll software products from PSDs may also produce government forms, but they do not offer direct tax filing services.

*Filing a client's payroll taxes directly with the State.* Employers make quarterly submissions of State payroll taxes using the State's QCR. Some PPFs use an employer's payroll to prepare the QCR and then submit the form and associated taxes directly to the State, assuming legal responsibility for the UI taxes. Direct filing of payroll taxes works well if the PPF includes employment counts on the QCRs and uses the correct definitions for these employment counts. If not, the effect of the resulting errors is a function of the size of the PPF's client base and the number of employees who work for those clients.

Sixty-two percent of PPFs file payroll taxes for at least some of their clients. The number of clients for whom they pay taxes ranges from 1 to greater than 100,000, with a median of around 100. This represents slightly more than half of the total number of client firms reported in this study. Of the PPFs that file payroll taxes, 15 percent do so for 90 percent or more of their clients. At the other end of the scale, around 20 percent of the PPFs file taxes for a small portion of their clients (10 percent or fewer). On average (median), PPFs that file payroll taxes do so for about 40 percent of their client firms.

*Preparing Forms that Contain Employment Counts.* Whether or not a PPF files its clients' payroll taxes, many PPF and commercial software products have the capability to generate the QCR and the CES form. The data show striking differences between PPFs and software packages, especially for the QCR. While 86 percent of PPFs prepare QCRs for their clients, only 61 percent of PSDs offer the capability in their payroll packages. Additionally, software products are nearly twice as likely as PPFs to produce the CES forms, although the percentage in both cases is much lower than that for the QCR.

As pointed out earlier, making a service or capability available does not ensure that clients use it. If a PPF prepares the QCR, we asked whether they did so for all clients, or if some clients completed the form themselves. (This is a moot point for PSDs, because the user decides whether or not to have the software generate the form.) Of the PPFs that file the QCR for some clients, nearly half (49 percent) do so for *all* their clients. Small PPFs were somewhat more likely than large PPFs to complete the QCR for all their clients (51 versus 39 percent).

**B. Reporting Picture/Error Profile**
1. Analysis of Measurement Errors

The consequences of measurement error in the three aspects of employment reporting (method, timing, and content) vary widely in severity. Seemingly incorrect reporting merely provides the potential for varying degrees of error in employment data. For example, if a PPF uses a "check count" method and employees always receive all wages (bonuses, overtime, etc.) on one check, the count is correct. The "worst" errors in record-keeping practices contain the most *potential* for measurement error. The actual extent of the error depends upon many factors, including individual record-keeping practices, client need for the data, State tax laws, Federal tax laws, and knowledge of the BLS definitions.

The questionnaire contained items asking about each of the three components of employment counts (method, timing, and content). The "method" question asked how employees were counted and offered these options: number of active employees, number of employees in personnel data file, count of time cards, count of checks, unduplicated count of individuals receiving checks, or some other method. The timing question determined the time period the employment count represented, e.g., current pay period, a specific day during the month, cumulative for the quarter, and so forth. The content question looked specifically at whether groups of employees were included or excluded from the employment count. If the respondent could not answer the question, the interviewer recorded "don't know," noting cases where the respondent volunteered that the decision was "up to the client."

The results for each of the components were analyzed separately for the PPFs and the PSDs to see

the impact each group has on employment counts. For the PPFs, the predominant finding is that 95 percent of the employment reported is determined with the correct method for counting employees. In addition, both the correct method and timing are used for 64 percent of the employment reported by PPFs. Results for the PSDs suggest that 37 percent of the firms use the correct method. As explained, the distribution of the correct method could be much higher since actual error is a function of the structure and content of client record-keeping systems. For PSDs, 9 percent utilize the correct time reference period. Keeping in mind the built-in flexibility of commercial payroll software, a combined 60 percent use either the current pay period, or leave the choice up to the client. These two responses could very well reflect the correct timing for the employment count, if used correctly by the client.

2. Payroll Processing Firms

*Method.* The correct method for the employment count is an unduplicated count of employees receiving pay. Of the PPFs responding to the question, 47 percent use this method. Results for all methods are given in Table 3. While the other methods do not adhere to the exact BLS definition, they should be fairly consistent over time. These methods might overstate actual employment, lose seasonality effects, or even be correct if they result in an unduplicated count of employees.

Nearly 95 percent of employment reported by PPFs is reported using the correct method. This suggests that the larger PPFs count employees in a way that is consistent with the BLS definition.

Table 3. PPF Employment Count Method

| Method | Percent of PPFs | Percent of Employment Reported by PPFs |
|---|---|---|
| Undup Count Employees | 47 | 95 |
| Active Employees | 12 | 0 |
| Number of Checks | 19 | 2 |
| Nr. Employee Records | 8 | 1 |
| Number of Time Cards | 2 | 0 |
| Don't Know | 9 | 1 |
| Other | 3 | 1 |
| Total | 100 | 100 |

*Timing.* The correct timing for the "all employment" count is the pay period including the 12th of the month. Thirty-one percent of the PPFs used this reference period. Results for all timing options appear in Table 4. As noted previously, the fact that some firms' responses do not adhere to the BLS definition does not necessarily mean that those firms produce incorrect counts. The current pay period will produce correct counts if the report is completed for the pay period of the 12th. This count, along with the count for one day during the month and the end of the quarter count, should produce fairly small errors which are consistent over time. The two cumulative counts, cumulative for month and any time in quarter/cumulative for quarter, would overstate actual employment to some degree.

Two-thirds of employment reported by PPFs (66 percent) is reported using the pay period of the 12th. Including the 1 percent that use the current pay period brings the total to 67 percent.

Table 4. PPF Employment Count Timing

| Timing | Percent of PPFs | Percent of Employment Reported by PPFs |
|---|---|---|
| Pay Period Including the 12th | 31 | 66 |
| Current Pay Period | 27 | 1 |
| As of Any Day in the Month | 9 | 1 |
| Cumulative for the Month | 5 | 0 |
| Any Time During the Quarter | 2 | 0 |
| End of the Quarter | 9 | 1 |
| Unknown [a] | 14 | 31 |
| Other | 3 | 0 |
| Total | 100 | 100 |

[a] Includes timing up to the client

*Content.* The content component of the BLS employment definition is the least likely to affect the employment count. Overall, 24 percent of the PPFs perform all eight aspects of the content component correctly. If the responses "don't know" (DK) and "up to the client" (UC) are considered as potentially correct, an additional 24 percent of the PPFs include and exclude all categories of employees correctly.

Approximately 54 percent of employment covered by PPFs is reported using all content groups correctly. If the responses DK and UC are recognized as potentially correct, an additional 10 percent of employment is reported using the correct content.

To measure the content of employee counts, the questionnaire listed eight separate groups or situations. Respondents indicated whether the employment counts they produced included or excluded employees in each group or situation.

3. Payroll Software Developers

*Method.* Of the PSDs responding to this question, 37 percent based the count on an unduplicated count of employees receiving checks. Results for all methods appear in Table 5. While the other methods do not adhere to our exact definition, they should be fairly consistent over time. Again, these methods might overstate actual employment, lose seasonality effects or even be correct if they actually result in an unduplicated count of employees.

*Timing.* Results for all timing options appear in Table 6. An estimated 9 percent of PSDs use the pay period including the 12th of the month as the time

Table 5. PSD Employment Count Method

| Method | Percent of PSDs |
|---|---|
| Unduplicated Count of Employees | 37 |
| Active Employees | 28 |
| Number of Checks | 6 |
| Number of Employee Records | 18 |
| Number of Time Cards | 1 |
| Don't Know | 0 |
| Other | 10 |
| Total | 100 |

reference period for counting employees. A larger percentage of PSDs (26 percent) use the current pay period as their timing option. This is correct if the report is prepared for the pay period including the 12th of the month. As stated earlier, this count, along with the one day during the month and the end of the quarter count, should produce fairly small errors which are consistent over time.

One another large group of PSDs (31 percent) leaves the timing aspect up to the client, which may also lead to correct results if the client knows what pay period to choose. All three cumulative counts tend to overstate actual employment.

Table 6. PSD Employment Count Timing

| Timing | Percent of PSDs |
|---|---|
| Pay Period Including the 12th | 9 |
| Current Pay Period | 26 |
| As of Any Day in the Month | 2 |
| Cumulative for the Month | 6 |
| Any Time During the Quarter | 3 |
| End of the Quarter | 1 |
| Cumulative for the Year | 1 |
| Up to the Client | 31 |
| Don't Know | 14 |
| Other | 7 |
| Total | 100 |

*Content*. An estimated 26 percent of PSDs handle each of the eight content issues correctly. If all responses of "don't know" (DK) and "up to the client" (UC) are interpreted as being correct, an additional 25 percent of the PSDs may count the appropriate categories of employees.

One big difference between PPFs and PSDs is the number of content items that PSDs leave to the user to implement, or about which respondents had no information. Only two PPF content items were left up to the client (or the respondent didn't know how an item was treated) more than 10 percent of the time. This contrasts with PSDs, where the smallest percentage of firms leaving the decision to include or exclude a group up to the user was 15 percent,

and the comparable portion for four different employee groups was about a third of PSDs.

## V. Discussion

BLS statistical programs that depend on employment counts increasingly obtain information produced by or in conjunction with standardized payroll software. The software is produced by two components of the payroll processing industry, PPFs and PSDs. Deviations from BLS employment count definitions built into in this software can have a noticeable effect on reported employment data.

BLS conducted a RAS of firms in the payroll processing industry to assess the industry's employment coverage and to evaluate measurement error in data produced by the industry. Survey results show that, overall, PPFs and PSDs prepare employment counts consistent with the BLS definition of employment. In many cases, the correct employees will be counted regardless of deviations from those definitions. Also, content "errors" may reflect extremely rare situations (e.g., strikes) or a very small percentage of employees (e.g., on leave without pay), and are often subject to interpretation by the client or software user.

Where measurement error does exist, the effect of it depends greatly on the number of employees covered by the PPF or the software product. Based on the percent of employment covered by PPFs, 95 percent of employment counts are derived using the correct method, 66 percent are based on the correct timing, and 54 percent use all 8 content items properly. Comparable percentages of employment for software firm products are not available, in part because of user-specified options in the software.

In sum, the payroll processing industry has the potential to introduce measurement error into BLS employment count data. These errors are relatively minimal for clients of payroll processing firms, and may be somewhat greater for employers who purchase commercial payroll software products to generate their payrolls. In both cases, improvements and corrections are possible. The BLS has established direct contact with members of the industry that provide incorrect employment counts and will encourage these firms to modify their reporting practices in accordance with BLS definitions.

## REFERENCE

Werking, George S., Patricia M. Getz, and Richard L. Clayton. 1993. "New Approaches to Controlling Response Error in Establishment Universe and Sample Data." Presented at the Annual Meeting of the American Statistical Association, August 1993.